Virology 285, 270–277 (2001)
doi:10.1006/viro.2001.0938, available online at http://www.idealibrary.com on IDEAL®

A Plasmid of Phytoplasma Encodes a Unique Replication Protein Having Both Plasmid- and Virus-like Domains: Clue to Viral Ancestry or Result of Virus/Plasmid Recombination?

Kenro Oshima, Shigeyuki Kakizawa, Hisashi Nishigawa, Tsutomu Kuboyama, Shin-ichi Miyata, Masashi Ugaki, and Shigetou Namba[1]

Laboratory of Bioresource Technology, Graduate School of Frontier Sciences, The University of Tokyo, Bunkyo-ku, Tokyo 113-8657, Japan

Received January 10, 2001; returned to author for revision March 13, 2001; accepted March 29, 2001

The genomes of most prokaryotic and eukaryotic single-stranded (ss) DNA viruses, and some prokaryotic plasmids such as pLS1, commonly replicate via a rolling circle replication (RCR) strategy, and thus the viruses are hypothesized to have evolved from the plasmids, although evidence for this view is sparse. We have sequenced a circular plasmid of 3933 nt, pOYW, obtained from onion yellows phytoplasma (OY-W), a cell-wall-less, unculturable prokaryote that inhabits the cytoplasm of both plant and insect cells. pOYW contains five open reading frames (ORFs) on the same strand and apparently replicates by an RCR mechanism. Its *rep* gene (ORF5) encodes a unique protein, pOYW-Rep, with an unprecedented structure. The N-terminal region of pOYW-Rep has similarities to the RCR initiator protein (Rep) of pLS1 family plasmids but, unlike the Rep of other plasmids, its C-terminal region was unexpectedly similar to the helicase domain of the replication-associated proteins (Rap) of eukaryotic viruses, especially circoviruses (ssDNA viruses of vertebrates). The pOYW-Rep was specifically detected in OY-W-infected plant phloem cells, suggesting that it is a functional protein. We suggest that an ancestral phytoplasma plasmid pOYW may have acquired a helicase domain from host phytoplasmal DNA, entered the surrounding eukaryotic cytoplasm, and subsequently evolved into an ancestral eukaryotic ssDNA virus. Alternatively, a pOYW ancestor could have obtained the helicase domain by recombination with a virus: this would be the first example of recombination between plasmids and viruses.

INTRODUCTION

Numerous theories have been proposed regarding the origin of viruses (Sinkovics *et al.,* 1998; Holland and Domingo, 1998) but none has been clearly proven. It is believed that most extant eukaryotic viruses are derived from ancestors that themselves arose from simpler prokaryotic or eukaryotic episomal replicons (Sinkovics *et al.,* 1998). For example, retroviruses very likely originated from cellular retrotransposons (Temin, 1980).

Some plant, animal, and bacterial single-stranded (ss) DNA viruses replicate by a rolling circle replication (RCR) mechanism (Berns, 1990; Stenger *et al.,* 1991). Many multicopy plasmids isolated from gram-positive and -negative bacteria also replicate by the RCR mechanism (Gruss and Ehrlich, 1989; Baas, 1985). These viruses and plasmids have considerable similarities in the details of their replication mechanism. RCR involves the generation of a site-specific nick at the plasmid or viral leading-strand origin by plasmid- or virus-encoded RCR initiator proteins, followed by covalent extension of the free 3′-hydroxyl end by a DNA polymerase (Gruss and Ehrlich, 1989). Based on nucleotide and amino acid sequence similarity, an evolutionary relationship between these replicons has been suggested. For example, pT181 family plasmids (Khan, 1997) and M13-related bacteriophages have nucleotide sequence similarity at the leading-strand origin (del Solar *et al.,* 1998). A major difference between the RCR initiator proteins (Rep) of plasmids and the replication-associated protein (Rap) of eukaryotic ssDNA viruses is that plasmid Rep, like the RCR initiator protein of ϕX174-related bacteriophages, is composed of only the RCR initiator domain, while viral Rap has the RCR initiator domain plus a helicase domain. Koonin and Ilyina (1992) demonstrated that the RCR initiator domains of Rap from the ssDNA geminiviruses (plant viruses) contain motifs similar to those found in the analogous domains of Rep from the pLS1 and pUB110 plasmid families (Khan, 1997). Thus, it is hypothesized that the eukaryotic ssDNA viruses have originated from prokaryotic RCR-type plasmids.

Phytoplasmas are unique prokaryotes belonging to the class *Mollicutes.* They are characterized by their small genome size, the absence of cell walls, and the fact that they inhabit eukaryotic cytoplasm. These microorganisms can infect the phloem sieve elements of plants and are transmitted between plants by phloem-

Novel sequence information appears under GenBank Accession No. AB056858.

[1] To whom correspondence and reprint requests should be addressed at Laboratory of Bioresource Technology, Graduate School of Frontier Sciences, The University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-8657, Japan. Fax: 81-424-69-8786. E-mail: snamba@ims.u-tokyo.ac.jp.

0042-6822/01 $35.00

feeding vector insects (McCoy *et al.,* 1989; Kirkpatrick, 1992; Lee and Davis, 1992). In our previous work on a wild strain (OY-W) of onion yellows phytoplasma, a 3.5-kb DNA fragment, believed to be derived from an RCR plasmid, was cloned (Kuboyama *et al.,* 1998). Here we show that an intact 3.9-kb plasmid cloned from the same strain and having a sequence closely related to that of the 3.5-kb fragment encodes a unique protein that is possibly responsible for replication. This protein contains not only a plasmid-like RCR initiator domain, but also a virus-like helicase domain. The relevance of this finding to the evolutionary history of viruses and plasmids is discussed.

RESULTS

Cloning and sequencing of a plasmid DNA from the OY-W phytoplasma

In our previous work, a 3.5-kb extrachromosomal DNA fragment was cloned from OY-W phytoplasma and was concluded to be derived from an RCR-type plasmid (Kuboyama *et al.,* 1998). However, a cloning site interrupted the 3′ end of the *rep* gene and the stop codon was not found in the fragment. Therefore, it was believed that the clone does not contain the complete plasmid sequence. To isolate a full-length plasmid, shotgun libraries of DNA from plant tissues infected with OY-W phytoplasma were screened, and two clones were selected from each of the shotgun libraries constructed using *Hin*dIII (WH-105) or *Pvu*II (WP-38). Sequence analysis indicated that both clones bore the same DNA fragment of about 3.9 kb and contained a replication protein gene (*rep*) similar to those previously reported for other RCR-type plasmids found in bacteria (Fig. 1A) (Kuboyama *et al.,* 1998). Moreover, it was clear that these two clones, which were cut at different sites (Fig. 2), each represented the same complete circular DNA. This DNA is designated pOYW. The complete sequence of pOYW was 3933 nucleotides long, and five open reading frames (ORFs) encoding proteins larger than 10 kDa were identified. The restriction sites, and the locations of the ORFs, are shown in Fig. 2.

The plasmid rep gene encodes a protein containing plasmid Rep and viral Rap-hel domains

The N-terminal region (residues 1–192) of the deduced amino acid sequence of ORF5 (377 aa) was similar (14.5% identity) to the putative replication initiator protein (Rep) of plasmids belonging to the pLS1 family, which is one of five major families of RCR-type plasmids (Fig. 1A) (Khan, 1997; del Solar *et al.,* 1993). The pLS1 family includes pKMK1 from *Mycoplasma mycoides* subsp. mycoides (17.5% identity) (King and Dybvig, 1992), *Leuconostoc lactis* 533 plasmid pCI411 (15.6% identity) (Coffey *et al.,* 1994), and *Streptococcus pneumoniae* plasmid pE194 (16.3% identity) (Fig. 1A) (Villafane *et al.,* 1987; Horinouchi and Weisblum, 1982). In the Rep proteins of these plasmids, five conserved motifs, RI to RV, have been identified previously (del Solar *et al.,* 1993). The Rep protein of pOYW was found to contain four of these five motifs, but one motif, RI, was absent (Figs. 1A and 1B).

Although the N-terminal region of the protein encoded by ORF5 showed considerable homology with Rep of pLS1 family plasmids, it showed much lower homology with Rep of plasmids belonging to four other RCR-type plasmid families, namely, pT181, pC194, pSN2, and pIJ101 (Khan, 1997). The N-terminal region also showed low (7.0%) homology with an RCR initiator protein of bacteriophage ϕX174.

The Walker A- and B-motifs (Saraste *et al.,* 1990), which are known as nucleotide-binding motifs, were identified in the approximately 100 amino acid residues in the pOYW-Rep protein C-terminal region (residues 193–302). Both of the characteristic consensus motifs, GXXXXGKS (Walker A-motif) and DD (Walker B-motif), were present (Fig. 1A). Neither of these motifs is reported in any replication protein from the RCR-type plasmids, including those of the pLS1 family. Furthermore, the region including these Walker A- and B-motifs (residues 204–247) showed great similarity to the replication-related protein (Rap) from eukaryotic small DNA viruses or some RNA viruses. Especially, it was most similar to several proteins of vertebrate viruses, the putative Rap from circoviruses [*Porcine circovirus* (Meehan *et al.,* 1997) (25.4% identity); *Beak and feather disease virus* (Niagro *et al.,* 1998) (26.3% identity)], and caliciviruses [*Human calicivirus* (Xi *et al.,* 1990) (18.5% identity); *Feline calicivirus* (Guiver *et al.,* 1992) (19.4% identity)] (Fig. 5). These viral regions are all known as helicase domains and are conserved in several other groups of plant and animal viruses with small ss- or dsDNA (papova-, parvo-, and geminiviruses) or RNA (picorna-, como-, and nepoviruses) genomes (Koonin and Ilyina, 1992). Such domains belong to the superfamily III helicase, which contains Walker A- and B-motifs (Gorbalenya *et al.,* 1990). Geminivirus Rap (AL1) also possesses a helicase domain; however, the helicase region of pOYW-Rep does not have much similarity to that of geminivirus-AL1 protein [*Beet curly top virus* (9.8% identity); *Tomato leaf curl virus* (9.2% identity)].

The deduced amino acid sequence of ORF4 (103 aa) shows similarity (36% identity) to the N-terminal two-thirds of the single-strand DNA binding (SSB) protein from *Bacillus subtilis* (Kuboyama *et al.,* 1998). The amino acid sequences encoded by pOYW ORFs 1, 2, and 3 did not show significant homology with previously reported proteins.

To confirm the molecular weight of the pOYW Rep protein (pOYW-Rep), Western blot analysis was performed with an anti-Rep protein antibody (Fig. 3C). The results indicated the existence of a protein with the

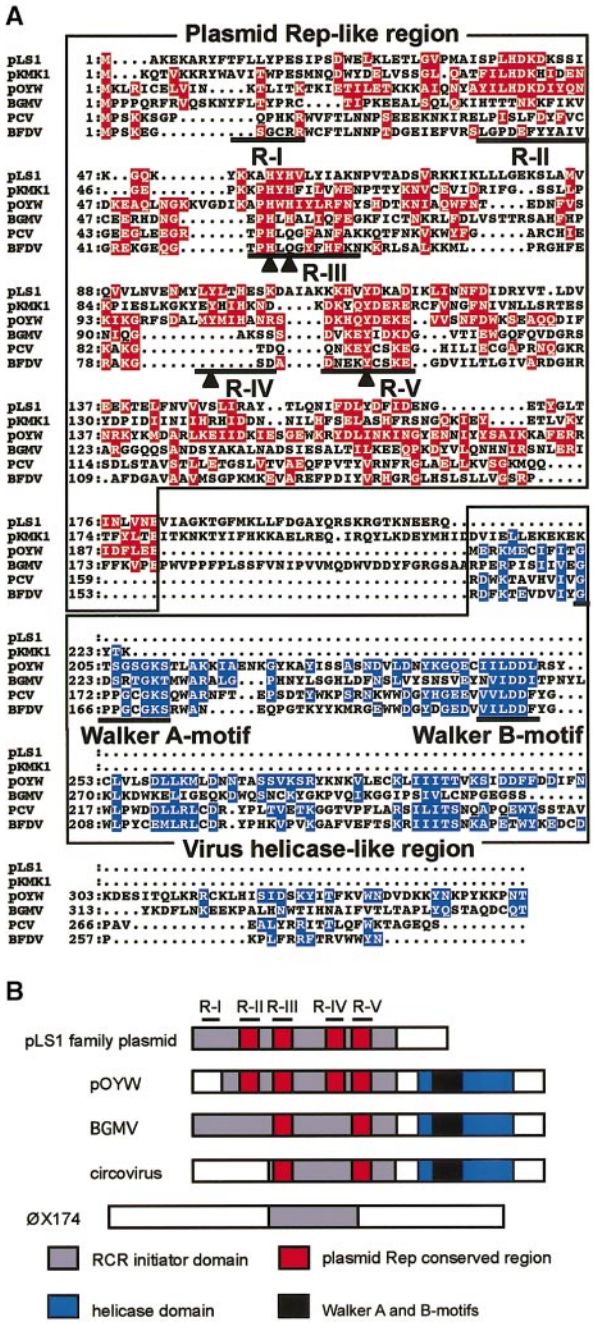

FIG. 1. (A) Comparisons of the deduced amino acid sequence of ORF5 with other Rep genes from RCR plasmids ("Plasmid Rep-like region," indicated as red) and of the C-terminal region of pOYW ORF5 and the replication-related proteins of circoviruses ("Virus helicase-like region," indicated as blue). Putative common regions of pLS1 family plasmids (R-I, R-II, R-III, R-IV, and R-V) described in the text are indicated. Arrowheads indicate the conserved H (domain R-III) and Y (domains R-IV and R-V) residues. Nucleotide binding motifs (Walker A- and B-motifs) are indicated. Amino acid sequences used for the alignment are the pOYW Rep protein and the replication-related proteins from geminivirus (BGMV, *Bean golden mosaic virus*) and circoviruses (PCV, *Porcine circovirus;* BFDV, *Beak* and *feather disease virus*). Residues in reverse contrast are identical or similar to pOYW-Rep residues; dots indicate gaps. (B) Domain organization of the pLS1 plasmid family, geminivirus, circovirus, bacteriophage ϕX174, and pOYW putative RCR initiator proteins. The conserved RCR initiator domain is shown in gray. The approximate positions of the four conserved motifs are shown.

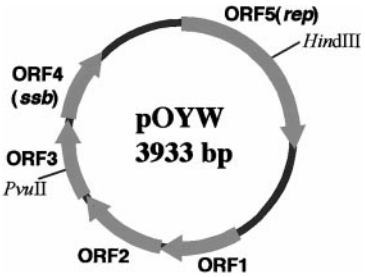

FIG. 2. Schematic diagram of pOYW. The positions of the five ORFs are shown. The largest ORF (ORF5) potentially encodes a replication-associated protein similar to the Rep proteins of geminiviruses, circoviruses, and other pLS1 plasmids.

expected molecular weight (44.6 kDa) in the soluble fraction. This confirmed that the Rep protein is expressed in the OY-W phytoplasma-infected plants. Although three other bands were detected besides the 45-kDa protein, it was considered that they were the background bands derived from a plant cell, because these bands were detected also in the extract from a healthy plant. Immunocytochemical treatment of thin sections observed using a light microscope showed that the Rep protein is characteristically detected in the phloem tissues and is colocalized to cells infected by OY phytoplasma as indicated by staining with antibodies to SecA (Fig. 4).

Southern and Northern hybridization analysis

Southern hybridization analysis was performed to confirm that pOYW existed in the phytoplasma as an extrachromosomal DNA molecule and that the full-length version of the plasmid had been cloned. A *rep* gene fragment was used as a probe against undigested total DNA extracted from an OY-W-infected plant (Fig. 3A, lane W). A major, slower-migrating band (indicated by an arrow-

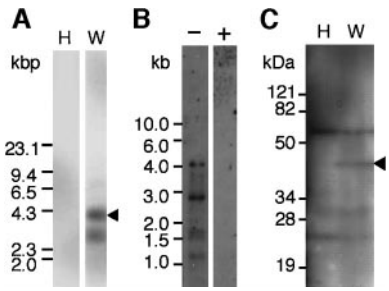

FIG. 3. Characterization of pOYW and Western blot analysis using an anti-Rep antibody. (A) Southern hybridization with the *rep* gene as a probe. Total DNA was extracted from a healthy plant (H) and an OY-W-infected plant (W). (B) Northern blot hybridization with plus- (+) and minus- (−) strand *rep* RNA used as probe. Total RNA was extracted from an OY-W-infected plant. (C) Western blot analysis using anti-Rep IgG. Extracts from the healthy (H) and OY-W-infected (W) plants were separated by 10% SDS–PAGE. Gels were blotted onto PVDF membranes and reacted with the anti-Rep IgG (1:1000 dilution) followed by the alkaline phosphatase-tagged secondary antibody (Amersham Pharmacia Biotech). Arrowhead in C indicates the phytoplasma-specific band with a molecular mass of ca. 45 kDa.

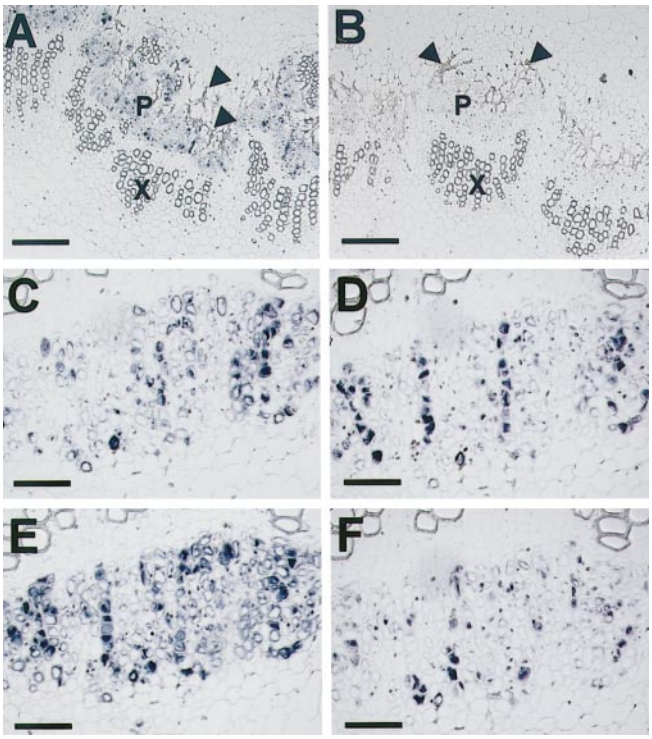

FIG. 4. Immunocytochemical detection of the pOYW-encoded Rep protein and the phytoplasma genome-encoded SecA protein in a vascular bundle of the stem of garland chrysanthemum plants. (A, B) Cross-sections of an OY-W-infected plant (A) and a healthy plant (B) reacted with an anti-Rep antibody (bars = 100 μm) showing phloem (P)-specific localization of the Rep protein. (C–F) Sequential cross-sections of an OY-W-infected plant reacted alternatively with an anti-SecA antibody (C, E) and with an anti-Rep antibody (D, F) (bars = 10 μm) showing colocalization of the two proteins. Goat anti-rabbit secondary antibody conjugated to alkaline phosphatase (AP) was used for detection. Localization of SecA and Rep protein was visualized as blue pigmentation. P, phloem; X, xylem. Arrowheads show areas of necrosis. Tissues were fixed in 4% (w/v) paraformaldehyde (Sass, 1958) and embedded in Paraplast Plus (Sherwood Medical, St. Louis, MO). Microtome sections (10 μm thick) were applied to charged slide glasses (Fisher Scientific, Pittsburgh, PA). The sections were deparaffinized in xylene and washed in PBS buffer. They were incubated in blocking buffer (10% FCS and 0.02% NaN_3 in PBS buffer) for 30 min at room temperature and incubated with a 1:1000 dilution of the anti-Rep or anti-SecA IgG in blocking buffer at 4°C overnight. The tissue sections were then washed in PBS-T buffer and incubated in a 1:1000 dilution of AP-coupled goat anti-rabbit IgG secondary antibody (Vector Laboratories, Inc., Burlingame, CA). For color development, sections were incubated with 0.34 mg/ml nitroblue tetrazolium salt and 0.175 mg/ml 5-bromo-4-chloro-3-indolyl phosphate toluidinium salt. After color development, the reaction was stopped by immersing the slides in TE (10 mM Tris–HCl, 1 mM EDTA), pH 7.5. The sections were dehydrated through a graded ethanol series and then mounted in Eukitt (O. Kindler, Germany).

head) and a minor band migrating faster were detected. Judging from their relative mobility, we presume that the major and minor bands were open and closed circular forms, respectively, of the 3.9-kb pOYW plasmid. No bands were detected with DNA extracted from a healthy plant (Fig. 3A, lane H). This result indicated that the cloned fragments were from the OY-W phytoplasma and that the plasmid was in fact full length.

To investigate whether pOYW is transcribed in a phytoplasma cell, Northern hybridization analysis was done using total RNA extracted from phytoplasma-infected plants and a *rep* gene fragment as a probe. When a minus strand of the *rep* gene was used as a probe to detect plus-strand RNA molecules (Fig. 3B, left), one major and several minor bands were detected from RNA extracted from an OY-W-infected plant. The size of the major band is about 3 kb and roughly corresponds to a size covering ORF5 through ORF4. Three base pairs after the termination codon of ORF4, there is a short stem–loop structure of 5′-AAGCCCCTAAACTAATAGGGGCTT-3′ (stem region is underlined), which may be related to transcription termination of this message. Other minor bands may be breakdown products or may represent minor initiation or termination events. No band was observed with the plus-strand RNA probe detecting minus-strand RNA molecules (Fig. 3B, right). These results indicate that the *rep* gene on pOYW is directionally transcribed in phytoplasma cells.

DISCUSSION

There are several lines of evidence to show that the pOYW is a plasmid-inhabiting phytoplasma. First, it has been detected only from phytoplasma-infected plants and never from healthy ones, and thus it is phytoplasma-specific (Kuboyama *et al.,* 1998). Second, no virus-like particles have been observed from the phytoplasma-infected pOYW-containing plants (Nishigawa *et al.,* 2001), and thus the pOYW is not likely to be a virus. Third, each of the five ORFs on pOYW has a Shine–Dalgarno (SD) sequence, which is indispensable for bacterial translation initiation, and thus can be translated within phytoplasma. Fourth, histochemical analysis using antibodies specific for Rep, encoded in pOYW, and SecA, encoded on phytoplasmal genomic DNA, indicated that these two proteins always colocalize. Fifth, Southern blot hybridization analysis showed that, in the phytoplasma genomes, there are no DNA sequences homologous to the pOYW sequence, and thus pOYW has not been integrated into the phytoplasma genome as a temperate phage would be.

Northern blot (Fig. 3B) and Western blot (Fig. 3C) analyses showed expression of the pOYW-Rep gene in the phloem tissues of OY-W-infected plants. Moreover, two plasmids closely related to pOYW, obtained from different OY-W strains, each contained the pOYW-*rep* gene, although one of the plasmids lacked two other ORFs (data not shown). These data suggest that the pOYW-*rep* gene is functional. This is the first evidence for expression of a phytoplasmal plasmid-encoded protein in phytoplasma-infected cells. As reported previously with the pOYW1 (a partial clone of pOYW) (Kuboyama *et al.,* 1998),

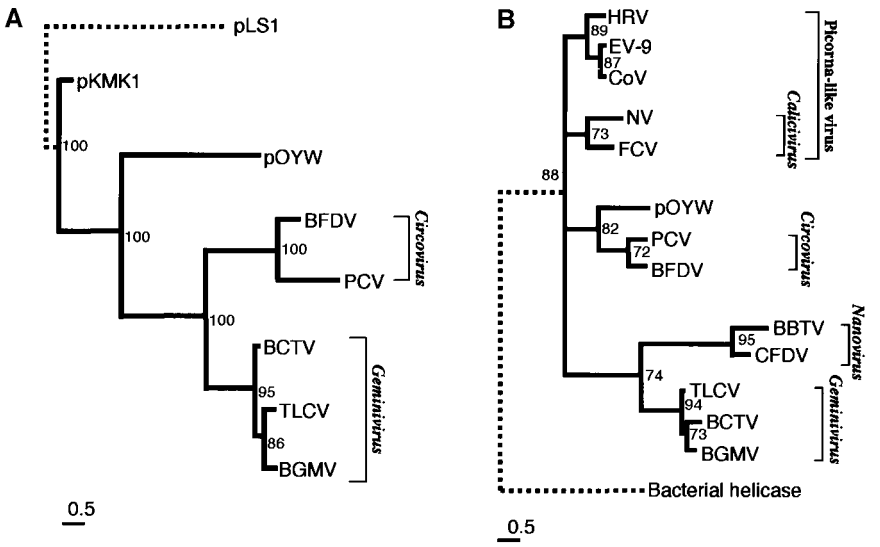

FIG. 5. (A) A maximum likelihood tree for the amino acid sequences of the N-terminal region of the pOYW Rep protein (amino acid position 1–192, see Fig. 1) and the corresponding regions of plasmid-encoded Rep and virus-encoded Rap proteins. Amino acid sequences of proteins that showed similarity with each ORF-encoded product were aligned by Clustal W version 1.6 (Thompson *et al.,* 1994). Maximum likelihood trees were inferred from aligned amino acid sequences by quartet puzzling using PUZZLE version 4, after positions including gaps had been excluded. Likelihoods were calculated by using the BLOSUM 62 substitution matrix. The sequence from pLS1-Rep was used as an outgroup to root the tree. The genetic distances of the amino acid sequences from each other were estimated by the amino acid substitution rate (*K* value). (B) A maximum likelihood tree for the amino acid sequences of the C-terminal region of pOYW-Rep protein (amino acid position 193–302, see Fig. 1) and the corresponding regions of the picorna-like virus 2C proteins and the circo-, nano-, and geminivirus Rap proteins. Sequences: picorna-like viruses (HRV, *Human rhinovirus* type 14; CoV, *Human coxsackievirus;* EV-9, *Echovirus* type 9), caliciviruses (FCV, *Feline calicivirus;* NV, *Human calicivirus* Norwalk virus serotype), nanoviruses (CFDV, *Coconut foliar decay virus;* BBTV, *Banana bunchy top virus*), and geminiviruses (BCTV, *Beet curly top virus;* TLCV, *Tomato leaf curl virus*). Other abbreviations are as described in the legend to Fig. 1.

the RCR of pOYW was confirmed by detecting the ssDNA replication intermediate that is specifically produced by the RCR mechanism (data not shown).

The N-terminal region of pOYW-Rep has extensive homology with the RCR initiator protein of the prokaryotic pLS1 plasmid family (del Solar *et al.,* 1993). The region has much less similarity to the RCR initiator proteins of plasmids belonging to four other RCR plasmid families, namely, pT181, pC194, pSN2, and pIJ101. Interestingly, the estimated molecular weight of the pOYW-Rep was 44.6 kDa and was ca. 15 kDa larger than that of Rep from the typical members of the pLS1 family (Khan, 1997). The extra size of pOYW-Rep is accounted for by an additional ca. 100-amino-acid domain at the C-terminal region that resembles helicase domains of Rap proteins of viruses which infect eukaryotes. Rep of pOYW does not resemble that of any RCR-type bacteriophages. For instance, while the RCR initiator proteins of pLS1 family plasmids are typically 250 amino acid residues in length and have the RCR initiator domain in the N-terminal region of the protein, that of well-studied bacteriophage ϕX174 is ca. 500 amino acids in length and has an RCR domain in the middle of the polypeptide (Koonin and Ilyina, 1993, see Fig. 6 in this reference).

When a maximum likelihood tree was made by analyzing the N-terminal region of pOYW-Rep, and of the Rap proteins of circoviruses and geminiviruses, with the N-terminal region of pLS1-Rep as an outgroup (Fig. 5A), pOYW-Rep is in the branch that includes circoviruses and geminiviruses. However, the evolutionary distances between pOYW-Rep and plasmid pKMK1-Rep and between pOYW-Rep and viral Rap are almost the same.

Interestingly, nucleotide sequences found at the nick site at the origin of replication of either the pLS1 family plasmids, circoviruses, or nanoviruses (Koonin and Ilyina, 1993; Mankertz *et al.,* 1998) are not found in pOYW. The equivalent sequence in pOYW is not yet identified for certain, although there are candidate sequences within the stem–loop-rich sequence between ORF4 and ORF5. An associated difference between pOYW-Rep and pLS1-Rep is that pOYW-Rep does not contain the R-I region, which is conserved in the pLS1 plasmids. It is speculated that the R-I motif is involved in the recognition of sequences surrounding the nick site and/or in the recognition of secondary structures in the nick region (del Solar *et al.,* 1993). Therefore, if the pOYW nick site is unique, so also might be a sequence in pOYW-Rep corresponding to the R-I motif.

As shown in Fig. 5B, a phylogenetic tree representing the Rep C-terminal (helicase) region revealed that pOYW-Rep was included in the circovirus Rap cluster. This region also showed a similarity to replication-associated RNA-binding proteins (2C proteins) encoded by picorna-like viruses, a large group of ssRNA viruses. The C-terminal helicase region of the circovirus Rap protein also has similarities to the 2C proteins (Gibbs and

Weiller, 1999). There have been no previous reports of plasmids encoding a replication-related protein with both plasmid- and virus-like domains. pOYW-Rep and its counterparts from pOYW-related plasmids (data not shown) thus constitute a novel class of prokaryotic plasmid-encoded protein.

It has been speculated that eukaryotic ssDNA viruses may have originated from a group of prokaryotic RCR plasmids, since they share the RCR strategy. Koonin and Ilyina (1992) showed that motifs in the RCR initiator domain are conserved between geminivirus Rap and pLS1-type Rep proteins, indicating an evolutionary relationship. Rap from present-day ssDNA viruses is invariably composed of two domains, the RCR initiator domain and the helicase domain, while Rep proteins from all known RCR plasmids contain only the RCR initiator domain. Thus, pOYW-Rep, possessing both the RCR initiator domain homologous to plasmid Rep and the helicase domain homologous to viral Rap, could be related to a viral Rap progenitor.

The initiation of pLS1-type plasmid replication requires host replication proteins such as a DNA helicase, a single-stranded DNA binding protein, and a DNA polymerase (Khan, 1997). However, pOYW encodes an SSB protein (ORF4) and its Rep (ORF5) has a helicase domain. Such a gene organization suggests that the only host factor that is required for pOYW replication is a DNA polymerase. Phytoplasmas have been reported to have genomes of only 0.6–2 Mb, among the smallest known for a prokaryotic organism (Marcone *et al.,* 1999). Phylogenetic analyses using 16S rRNA gene sequences (Namba *et al.,* 1993; Lee *et al.,* 2000) and other genes (Namba *et al.,* unpublished data) indicate that the phytoplasma diverged from the *Acholeplasma* branch of the *Mollicutes* phylogenetic tree. The *Mollicutes* branch diverged from the free-living *Streptococcus* branch of the gram-positive phylogenetic tree. This suggests that the ancestor of phytoplasmas was a free-living bacterium, probably with a larger genome. During adaptation to the cytoplasm of eukaryotic cells, phytoplasmas may have become more capable of utilizing the cellular functions of host cells and less dependent on their own genes and so perhaps have evolved toward smaller genome size. In this situation, we speculate that the plasmids in the ancestral phytoplasmas evolved in the reverse direction; their evolution could have involved acquiring and using more phytoplasma genes for their own DNA replication. This evolutionary view is supported by the fact that pOYW encodes not only the RCR initiation function, but also the helicase and SSB functions. Living within eukaryotic cytoplasm probably led to the loss of the ancestral phytoplasma's cell wall, presumably in order to facilitate the exchange of substances with the surrounding cytoplasm. We speculate that the enhanced exchange of substances between prokaryotic and eukaryotic cytoplasms could have facilitated the entry of prokaryotic plasmids into eukaryotic cytoplasm.

Phytoplasmas replicate in both plant and insect (leafhopper) cells. Although they may induce severe phytopathological symptoms in plants, they do not induce any adverse effects in the leafhoppers, suggesting that the association of phytoplasmas and insects is more ancient than that with plants. ssDNA viruses encoding Rap proteins closely related to pOYW-Rep, i.e., circoviruses, geminiviruses, and nanoviruses, infect vertebrates or plants, and some are transmitted by insects. These facts suggest a possible evolutionary relationship between phytoplasmal plasmids and ssDNA viruses.

Taking the above discussions into consideration, if ssDNA viruses did evolve from a prokaryotic plasmid, pOYW may be the first report of a descendant of this plasmid. However, other hypotheses cannot be excluded. One is that these two replicons are evolutionarily independent and do not share any ancestors. Alternatively, the chimeric nature of pOYW-Rep could be the result of recombination between a phytoplasmal plasmid and circovirus DNA. Indeed, some instances of recombination between different viral replication-associated protein genes have been reported (Gibbs and Weiller, 1999; Saunders and Stanley, 1999). Recombination provides the viruses with the potential for rapid adaptation to new hosts and environmental conditions, and it plays a major evolutionary role by creating genetic diversity within the virus population (Domingo and Holland, 1997). Interspecific recombination has been well documented in many virus families (Roossinck, 1997), for example, between RNA viruses (White *et al.,* 1995), between DNA viruses (Zhou *et al.,* 1997), and between RNA and DNA viruses (Morse *et al.,* 1992). However, no evidence of recombination between eukaryotic virus genomes and prokaryotic plasmids has been reported. Therefore, if pOYW is the result of a plasmid–virus recombination event, it is the first such example. We hope our work will encourage further investigations into possible links in the evolutionary histories of viruses and plasmids.

MATERIALS AND METHODS

Extraction, cloning, sequencing, and sequence analysis of phytoplasma DNA

A wild strain of onion yellows phytoplasma (OY-W), isolated in Saga Prefecture, Japan, in May 1982 (Shiomi *et al.,* 1996) was used. To construct a DNA library, DNA was extracted from phytoplasma-enriched fractions prepared from OY-W-infected garland chrysanthemums (*Chrysanthemum coronarium*) propagated by serial transmission of OY-W by the leafhopper vector (*Macrosteles striifrons*) as previously described (Kuboyama *et al.,* 1998; Lee *et al.,* 1988). Control preparations were obtained from OY-W-free plants. For Southern hybridization, the DNA from OY-W-diseased and healthy plants was

isolated from these preparations as described previously (Kuske and Kirkpatrick, 1990). The DNA isolated from the OY-W-enriched fraction was completely digested with either *Hin*dIII or *Pvu*II, ligated into pUC18, and used to transform *Escherichia coli* JM109. The DNA library obtained was then used for differential screening. Probes were prepared by *Hin*dIII digestion of DNA isolated from either OY-W-infected or healthy plants. Clones that hybridized only with the probe prepared from OY-W-infected plants were identified.

Each DNA insert was sequenced by cycle sequencing using *Taq* FS DNA polymerase and fluorescent-dideoxy terminators (Applied Biosystems, Foster City, CA). The DNA fragments were separated by electrophoresis and analyzed using an automated Applied Biosystems 377 DNA sequencer. The degree of similarity of all five ORFs to other known sequences was ascertained by computer-assisted homology analyses using the World Wide Web (http://www.genome.ad.jp/) sequence interpretation tools (Institute of Medical Science, The University of Tokyo, Tokyo, Japan) and the BLAST algorithm (Altschul *et al.,* 1990). Maximum likelihood trees were constructed as described in the legend to Fig. 5.

Preparation of a polyclonal antibody against the pOYW-Rep protein

An expression plasmid carrying a histidine-tagged Rep fusion protein was constructed using the pET system (Novagen, Madison, WI) by amplifying a complete *rep* gene fragment (1131 bp) with primers Rep-N (5′-AGG ATC CCA TAT GAA ATT AAG AAT CTG CGA ACT TG-3′) and Rep-C (5′-AGA GCT CGA GTA TTA AAG ATG TTA GAT AAT AAT ACT-3′). The PCR product was digested with *Bam*HI and *Xho*I and inserted into pET30a. The fusion protein was expressed in *E. coli* and applied to a NTA nickel column; the column was then washed with TBS buffer and the fusion protein eluted with TBS buffer containing 10 mM imidazole. One milligram of purified protein suspended in 1 ml of water and emulsified with 1.1 vol of Freund's complete adjuvant (DIFCO, Detroit, MI) was injected intramuscularly into a New Zealand white rabbit. After 2 weeks, a booster injection containing 1 mg protein was given; serum was obtained after 4 weeks. The IgG fraction was purified from the serum using a protein A column (Bio-Rad, Hercules, CA).

Polyacrylamide gel electrophoresis and Western blotting

Crude protein extracts were prepared from both a healthy plant and an OY-W phytoplasma-infected plant by the method used to obtain the phytoplasma-enriched fraction for DNA isolation. Sodium dodecyl sulfate (SDS)–polyacrylamide gel electrophoresis was conducted as described by Laemmli (1970). For immunoblot analysis (Villafane *et al.,* 1987), gels were blotted onto polyvinylidene difluoride (PVDF) membranes (Millipore, Bedford, MA) with an electroblotting apparatus (Bio-Rad). The blots were developed with AlkPhos detection reagent (Amersham Pharmacia Biotech, UK) according to the supplier's instructions.

Southern blot analysis

DNA purified from the phytoplasma-enriched fraction and from control preparations was fractionated by agarose gel electrophoresis, transferred onto a nylon membrane, and probed with a DIG-UTP-labeled 700-bp *rep* gene fragment which was PCR-amplified by rep-1 (5′-TAT TTA TCA AAA TGA TAA AGA GGC TC-3′) and rep-2 (5′-CAA CGA CGT TTT AAT TGA GTA ATA C-3′) primers. Hybridization and detection were carried out using a DIG luminescent detection kit (Boehringer-Mannheim Biochemicals, Mannheim, Germany) according to the manufacturer's instructions.

Northern blot analysis

Total RNA was extracted from healthy, OY-W-infected garland chrysanthemum tissue, separated by agarose gel electrophoresis, and transferred to a nylon membrane (Amersham Pharmacia Biotech) (Sambrook *et al.,* 1989). Riboprobes were labeled with DIG-UTP by transcribing the subcloned *rep* gene using the T7 and SP6 RNA polymerases (Boehringer-Mannheim Biochemicals). Hybridization and detection were carried out using a DIG luminescent detection kit according to the manufacturer's instructions.

Immunocytochemical analysis

To analyze the localization of the plasmid-encoded Rep in plant tissues, immunocytochemical analysis of thin tissue sections was performed using anti-Rep IgG and a light microscope. The distribution of phytoplasma in the tissue was determined similarly, but with the aid of antiserum to a phytoplasma surface protein, SecA (unpublished data).

ACKNOWLEDGMENTS

We greatly thank Dr. Bryan Harrison, Scottish Crop Research Institute, United Kingdom, for helpful discussions and critical reading of the manuscript. This work was supported partly by Grants-in-Aid from the Ministry of Education, Science, and Culture of Japan (09460155) and by the Program for Promotion of Basic Research Activities for Innovative Biosciences (PROBRAIN) of the Ministry of Agriculture, Forestry, and Fishers of Japan.

REFERENCES

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* **215,** 403–410.

Baas, P. D. (1985). DNA replication of single-stranded *Escherichia coli* DNA phages. *Biochim. Biophys. Acta* **825,** 111–139.

Berns, K. I. (1990). Parvovirus replication. *Microbiol. Rev.* **54,** 316–329.

Coffey, A., Harrington, A., Kearney, K., Daly, C., and Fitzgerald, G. (1994).

Nucleotide sequence and structural organization of the small, broad-host-range plasmid pCI411 from *Leuconostoc lactis* 533. *Microbiology* **140,** 2263–2269.

del Solar, G., Giraldo, R., Ruiz-Echevarria, M. J., Espinosa, M., and Diaz-Orejas, R. (1998). Replication and control of circular bacterial plasmids. *Microbiol. Mol. Biol. Rev.* **62,** 434–464.

del Solar, G., Moscoso, M., and Espinosa, M. (1993). Rolling circle-replicating plasmids from gram-positive and gram-negative bacteria: A wall falls. *Mol. Microbiol.* **8,** 789–796.

Domingo, E., and Holland, J. J. (1997). RNA virus mutations and fitness for survival. *Annu. Rev. Microbiol.* **51,** 151–178.

Gibbs, M. J., and Weiller, G. F. (1999). Evidence that a plant virus switched hosts to infect a vertebrate and then recombined with a vertebrate-infecting virus. *Proc. Natl. Acad. Sci. USA* **96,** 8022–8027.

Gorbalenya, A. E., Koonin, E. V., and Wolf, Y. I. (1990). A new superfamily of putative NTP-binding domains encoded by genomes of small DNA and RNA viruses. *FEBS Lett.* **262,** 145–148.

Gruss, A., and Ehrlich, S. D. (1989). The family of highly interrelated single-stranded deoxyribonucleic acid plasmids. *Microbiol. Rev.* **53,** 231–241.

Guiver, M., Littler, E., Caul, E. O., and Fox, A. J. (1992). The cloning, sequencing and expression of a major antigenic region from the feline calicivirus capsid protein. *J. Gen. Virol.* **73,** 2429–2433.

Holland, J., and Domingo, E. (1998). Origin and evolution of viruses. *Virus Genes* **16,** 13–21.

Horinouchi, S., and Weisblum, B. (1982). Nucleotide sequence and functional map of pE194, a plasmid that specifies inducible resistance to macrolide, lincosamide, and streptogramin type B antibodies. *J. Bacteriol.* **150,** 804–814.

Khan, S. A. (1997). Rolling-circle replication of bacterial plasmids. *Microbiol. Mol. Biol. Rev.* **61,** 442–455.

King, K. W., and Dybvig, K. (1992). Nucleotide sequence of *Mycoplasma mycoides* subspecies mycoides plasmid pKMK1. *Plasmid* **28,** 86–91.

Kirkpatrick, B. C. (1992). Mycoplasma-like organisms: Plant and invertebrate pathogens. *In* "The Prokaryotes" (A. Balows, H. G. Truper, M. Dworkin, W. Harder, and K. H. Schleifer, Eds.), pp. 4050–4067. Springer-Verlag, New York.

Koonin, E. V., and Ilyina, T. V. (1992). Geminivirus replication proteins are related to prokaryotic plasmid rolling circle DNA replication initiator proteins. *J. Gen. Virol.* **73,** 2763–2766.

Koonin, E. V., and Ilyina, T. V. (1993). Computer-assisted dissection of rolling circle DNA replication. *BioSystems* **30,** 241–268.

Kuboyama, T., Huang, C. C., Lu, X., Sawayanagi, T., Kanazawa, T., Kagami, T., Matsuda, I., Tsuchizaki, T., and Namba, S. (1998). A plasmid isolated from phytopathogenic onion yellows phytoplasma and its heterogeneity in the pathogenic phytoplasma mutant. *Mol. Plant Microbe Interact.* **11,** 1031–1037.

Kuske, C. R., and Kirkpatrick, B. C. (1990). Identification and characterization of plasmids from the western aster yellows mycoplasmalike organism. *J. Bacteriol.* **172,** 1628–1633.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature (London)* **227,** 680–685.

Lee, I. M., and Davis, R. E. (1992). Mycoplasmas which infect plants and insects. *In* "Mycoplasmas: Molecular Biology and Pathogenesis" (R. N. McElhaney, L. R. Finch, J. B. Baseman, and J. Maniloff, Eds.), pp. 379–390. Am. Soc. Microbiol., Washington, DC.

Lee, I. M., Davis, R. E., and Gundersen-Rindal, D. E. (2000). Phytoplasma: Phytopathogenic mollicutes. *Annu. Rev. Microbiol.* **54,** 221–255.

Lee, I. M., Davis, R. E., Hammond, R., and Kirkpatrick, B. C. (1988). Cloned riboprobe for detection of a mycoplasmalike organism. *Biochem. Biophys. Res. Commun.* **155,** 443–448.

Mankertz, A., Mankertz, J., Wolf, K., and Buhk, H. J. (1998). Identification of a protein essential for replication of porcine circovirus. *J. Gen. Virol.* **79,** 381–384.

Marcone, C., Neimark, H., Ragozzino, A., Lauer, U., and Seemüller, E. (1999). Chromosome sizes of phytoplasmas composing major phylogenetic groups and subgroups. *Phytopathology* **89,** 805–810.

McCoy, R. E., Caudwell, A., Chang, C. J., Chen, T. A., Chiykowski, L. N., Cousin, M. T., Dale, J. L., de Leeuw, G. T. N., Golino, D. A., Hackett, K. J., Kirkpatrick, B. J., Marwitz, R., Petzold, H., Sinha, R. C., Sugiura, M., Whitcomb, R. F., Yang, I. L., Zhu, B. M., and Seemüller, E. (1989). Plant diseases associated with mycoplasma-like organisms. *In* "The Mycoplasmas" (R. F. Whitcomb and J. G. Tully, Eds.), Vol. V, pp. 546–640. Academic Press, New York.

Meehan, B. M., Creelan, J. L., Mcnulty, M. S., and Todd, D. (1997). Sequence of porcine circovirus DNA: Affinities with plant circoviruses. *J. Gen. Virol.* **78,** 221–227.

Morse, M. A., Marriot, A. C., and Nuttall, P. A. (1992). The glycoprotein of Thogoto virus (a tick-borne orthomyxo-like virus) is related to the baculovirus glycoprotein GP64. *Virology* **186,** 640–646.

Namba, S., Oyaizu, H., Kato, S., Iwanami, S., and Tsuchizaki, T. (1993). Phylogenetic diversity of phytopathogenic mycoplasmalike organisms. *Int. J. Syst. Bacteriol.* **43,** 461–467.

Niagro, F. D., Forsthoefel, A. N., Lawther, R. P., Kamalanathan, L., Ritchie, B. W., Latimer, K. S., and Lukert, P. D. (1998). Beak and feather disease virus and porcine circovirus genomes: Intermediates between the geminiviruses and plant circoviruses. *Arch. Virol.* **143,** 1723–1744.

Nishigawa, H., Miyata, S., Oshima, K., Sawayanagi, T., Komoto, A., Kuboyama, T., Matsuda, I., Tsuchizaki, T., and Namba, S. (2001). *In planta* expression of a protein encoded by the extrachromosomal DNA of a phytoplasma and related to geminivirus replication proteins. *Microbiology* **147,** 507–513.

Roossinck, M. J. (1997). Mechanisms of plant virus evolution. *Annu. Rev. Phytopathol.* **35,** 191–209.

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). *In* "Molecular Cloning: A Laboratory Manual," 2nd ed. Cold Spring Harbor Laboratory Press, New York.

Saraste, M., Sibbald, P. R., and Wittinghofer, A. (1990). The P-loop—A common motif in ATP- and GTP-binding proteins. *Trends Biochem. Sci.* **15,** 430–434.

Sass, A. E. (1958). *In* "Botanical Microtechnique," 3rd ed. Iowa State Univ. Press, Ames.

Saunders, K., and Stanley, J. (1999). A nanovirus-like DNA component associated with yellow vein disease of *Ageratum conyzoides:* Evidence for interfamilial recombination between plant DNA viruses. *Virology* **264,** 142–152.

Sinkovics, J., Horvath, J., and Horak, A. (1998). The origin and evolution of viruses. *Acta Microbiol. Immunol. Hung.* **45,** 349–390.

Shiomi, T., Tanaka, M., Wakiya, H., and Zenbayashi, R. (1996). Occurrence of Welsh onion yellows. *Ann. Phytopathol. Soc. Jpn.* **62,** 258–260.

Stenger, D. C., Revington, G. N., Stevenson, M. C., and Bisaro, D. M. (1991). Replicational release of geminivirus genomes from tandemly repeated copies: Evidence for rolling-circle replication of a plant viral DNA. *Proc. Natl. Acad. Sci. USA* **88,** 8029–8033.

Temin, H. M. (1980). Origin of retroviruses from cellular moveable genetic elements. *Cell* **21,** 599–600.

Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994). Clustal W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22,** 4673–4680.

Villafane, R., Bechhofer, D. H., Narayanan, C. S., and Dubnau, D. (1987). Replication control genes of plasmid pE194. *J. Bacteriol.* **169,** 4822–4829.

White, P. S., Morales, F., and Roossinck, M. J. (1995). Interspecific reassortment of genomic segments in the evolution of cucumoviruses. *Virology* **207,** 334–337.

Xi, J. N., Graham, D. Y., Wang, K. N., and Estes, M. K. (1990). Norwalk virus genome cloning and characterization. *Science* **14,** 1580–1583.

Zhou, X., Liu, Y., Calvert, L., Munoz, C., Otim-Nape, G. W., Robinson, D. J., and Harrison, B. D. (1997). Evidence that DNA-A of a geminivirus associated with severe cassava mosaic disease in Uganda has arisen by interspecific recombination. *J. Gen. Virol.* **78,** 2101–2111.